



## ANÁLISE E PREVISÃO DE SUÍNOS VIVOS ATRAVÉS DA REGRESSÃO LINEAR: UM ESTUDO DE CASO<sup>1</sup>

Kellyn Aretha Zotti Baron<sup>2</sup>, Isabela Fernanda Bonin<sup>2</sup>; Thaynara Lopes Faria<sup>2</sup>; Lais Caroline Cremon<sup>2</sup>; Edward Seabra Júnior<sup>3</sup>; Samuel Bellido Rodrigues<sup>4</sup>

<sup>1</sup>Apresentado no 2º Seminário de Engenharia de Energia na Agricultura: 30/11/2017-UNIOESTE, *Campus* Cascavel.

<sup>2</sup>Universidade Tecnológica Federal do Paraná - UTFPR – Bacharelado em Engenharia de Produção - Medianeira PR.

<sup>3</sup>Universidade Estadual do Oeste do Paraná – UNIOESTE, PPGEA – Programa de Pós-Graduação em Engenharia de Energia na Agricultura – Nível Mestrado, Cascavel-PR.

<sup>4</sup>Universidade Tecnológica Federal do Paraná - UTFPR – Professor - Medianeira PR

### RESUMO

Tendo em vista a evolução da suinocultura brasileira no cenário mundial, nota-se o aumento da competitividade e a redução da lucratividade dos produtores, tornando-se necessário o investimento no gerenciamento das granjas. Deste modo, é importante o uso da estatística para análise de dados através de métodos quantitativos para auxiliar na resolução de problemas e tomada de decisão. Este trabalho utilizou regressão linear simples para analisar os dados e prever o número de leitões nascidos vivos em uma empresa localizada na cidade de São Miguel do Iguaçu no Oeste do Paraná. Aplicando estas ferramentas estatísticas percebe-se que neste caso, o método pode ser utilizado para analisar e prever a quantidade de suínos nascidos vivos.

**Palavras-chave** Suinocultura; Regressão linear; Gestão de processo.

## ABSTRACT

Having in mind the evolution of the Brazilian swine breeding in the world stage, there is noticed the increase of the competitiveness and the reduction of the profitability of the producers, making investment in the management of farms necessary. In this way, the use of the statistic is important for data analysis through quantitative methods to help in the solve of problems and to make a decision. This study used the simple linear regression to analyse the data and to predict the number of lively born pigs in an enterprise located in São Miguel do Iguaçu in the West of the Paraná. Applying these statistical tools it is seen that in this case, the method can be used to analyse and to predict the quantity of lively born pigs.

**Keyword** Swine breeding; Linear regression; Process management.

## Introdução

A suinocultura vem crescendo muito nos últimos anos no Brasil. Segundo Ativos da suinocultura (2015), o país ocupa a quarta posição de produção e exportação deste tipo de carne no mundo. Em primeiro lugar encontra-se a China, seguido da União Européia e Estados Unidos.

Conforme Rohr e Costa (2016), a suinocultura é uma das atividades pecuárias que estão sob maior expansão no Brasil. Com este fato, as margens lucrativas dos produtores vêm decrescendo cada vez mais, por este motivo torna-se importante investir no gerenciamento das granjas suinícolas. Conforme a Associação Brasileira de Proteína Animal (2017), ocorreu um aumento de cerca de 68% na produção de carne suína entre os anos de 2000 e 2016.

Com isso, percebe-se a importância das ferramentas estatísticas para gerar analisar e prever os dados futuros. Estes métodos auxiliam o gestor na tomada de decisão (FERNANDES; GODINHO FILHO, 2010).

A utilização da regressão linear como método preditivo é muito importante dentro do processo produtivo, pois a partir dele, pode-se prever custos, gastos, lucratividade, produção, entre outros dados importantes para o processo de tomada de decisão.

Segundo Krajewski, Ritzman e Malhotra (2014), previsão da demanda é uma estimativa futura sobre o mercado. As empresas necessitam dimensionar suas demandas de modo que elas estejam alinhadas com a sua capacidade produtiva, pois desta forma seria possível prevenir as

perdas relacionadas com o material, mão-de-obra, energia e ausência de produtos para atendimento do mercado (TUBINO, 2007).

Conforme Corrar (2007), dentre as informações coletadas sobre o processo produtivo, pode-se encontrar algumas variáveis que auxiliam no aumento da capacidade produtiva. Deste modo, o modelo de regressão linear simples tem por objetivos prever valores destas variáveis.

## **Materiais e métodos**

O estudo foi realizado em uma granja do tipo UPL (Unidade Produtora de Leitões), que fica localizada na cidade de São Miguel do Iguaçu no Oeste do Paraná. Este tipo de granja funciona como base da cadeia produtiva de suínos, pois comporta reprodução de fêmeas e produção de leitões. O local possui cerca de um estoque de mil matrizes ativas.

Os dados utilizados deste trabalho foram cedidos pela empresa UPL. Estes, estão dispostos em forma de indicadores internos de controle de processo em uma planilha eletrônica. Foram utilizados dados do período de um ano, que são aferidos semanalmente, ou seja, os dados são de cinquenta e três semanas.

A empresa possui cinco setores, porém, para análise do processo produtivo em si, serão analisados somente os três setores da produção, que são: gestação, maternidade e creche.

O setor gestação é responsável pela inseminação que é realizada em média 3,4 vezes por fêmea em cada ciclo (aproximadamente 7 meses), dependendo do sucesso da inseminação. As fêmeas ficam neste local até perto da data de parição, e depois são levadas para a maternidade.

A maternidade recebe as fêmeas gestante do setor gestação e acompanha os partos. Neste local presta-se toda assistência necessária aos leitões e fêmeas até o desmame.

O setor creche recebe os leitões desmamados na maternidade, e acompanha-os até a venda final, prestando assistência alimentar e remédios.

Por meio de análises juntamente com o gerente da empresa, definiu-se as variáveis que influenciam na prática, na quantidade de leitões nascidos vivos, e estas estão descritas na Figura 1.

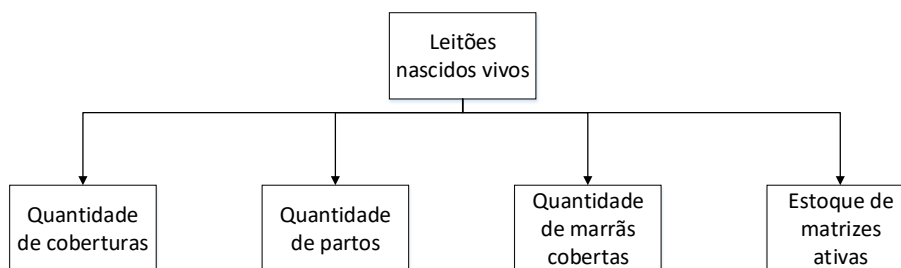


Figura 1- Fluxograma de variáveis analisadas

Conforme a Figura 1, a variável quantidade de coberturas mensura quantas fêmeas foram inseminadas. Quantidade de partos indica quantos partos ocorreram. Quantidade de marrãs cobertas indica quantas fêmeas jovens foram inseminadas. Por último estoque de matrizes ativas, mensura quantas fêmeas ativas existem no processo.

O modelo de regressão linear, segundo Krajewski, Ritzman e Malhotra (2014), é um dos modelos causais mais conhecidos e utilizados e define que a variável dependente é relacionada a uma variável independente por uma equação linear. Conforme Monks (1987), a regressão linear simples mostra correlação entre duas variáveis, ajustando uma reta entre elas.

A equação da regressão linear, conforme Martins e Laugen (2013) é dada por:

$$y = a + bx \quad (1)$$

Onde  $y$  representa a variável dependente,  $a$  o coeficiente de interseção,  $b$  o coeficiente de inclinação e  $x$  a variável independente.

Werkema (1996), mostra que a análise de regressão processa as informações contidas em um conjunto de dados de forma a gerar um modelo que mostre o relacionamento existente entre as variáveis de interesse de um processo, tendo como objetivos a descrição, predição, controle e a estimação.

## Resultados e discussão

Primeiramente foi feita a análise de tendência e sazonalidade da variável dependente. O resultado está apresentado na Figura 2.

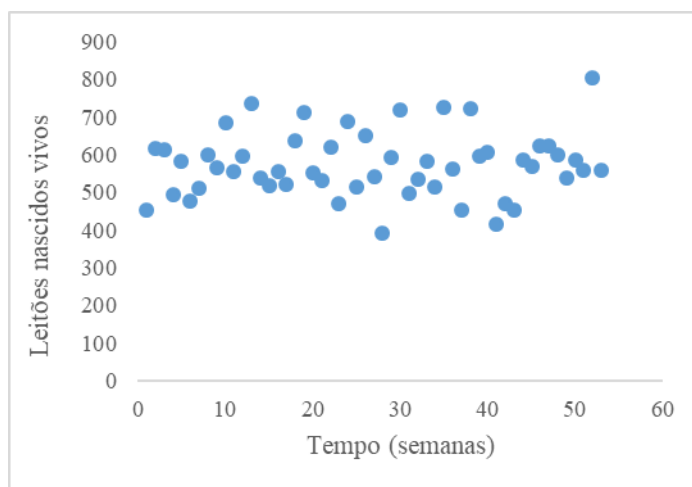


Figura 2- Dados de leitões nascidos vivos

Através da análise gráfica da Figura 2, observou-se que não há tendência e nem sazonalidade aparente, pois como pode-se visualizar, não existe tendência de crescimento ou decrescimento da série e também não há determinada época do ano em que os dados se repitam, pois estão dispersos e não seguem uma tendência de crescimento ou repetição.

Primeiramente, foi analisada a variável explicativa quantidade de partos, em relação à variável nascidos vivos. A Figura 3 mostra a dispersão dos dados de quantidade de partos em relação à quantidade de leitões nascidos vivos.

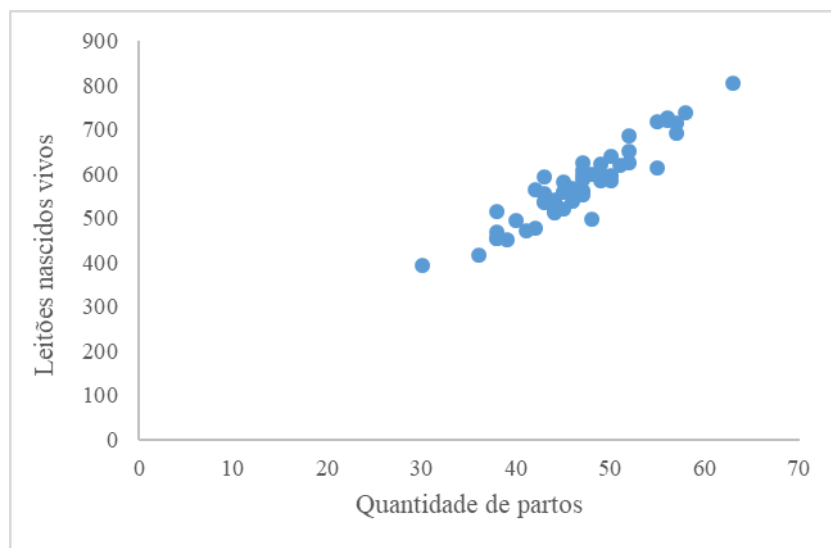


Figura 3- Quantidade de partos *versus* Nascidos Vivos

Por meio da Figura 3, percebeu-se que há uma certa tendência na variável, conforme o tempo passa em relação ao número de leitões nascidos vivos.

Após a análise do gráfico apresentado, foi feita regressão linear entre as variáveis explicativa quantidade de partos e a variável resposta nascidos vivos. Utilizou-se um nível de confiança de 95%. Os resultados podem ser visualizados na Tabela 1.

ESTATÍSTICA DA REGRESSÃO								
R múltiplo	0,94							
R-Quadrado	0,88							
R-quadrado ajustado	0,88							
Erro padrão	30,38							
Observações	53							
	<i>GL</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>			
Regressão	1	339503,81	339503,81	367,93	5,64E-25			
Resíduo	51	47059,51	922,74					
Total	52	386563,32						
	<i>Coefi- c entes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção quantidade de coberturas	-25,52	31,54	-0,81	0,42	-88,831	37,799	-88,831	37,799
	12,86	0,67	19,18	5,6e-25	11,52	14,21	11,52	14,21

Tabela 1 – Resultados da regressão considerando as variáveis de Quantidade de Partos e Nascidos Vivos

Na Tabela 1 percebe-se que a relação entre as variáveis é significativa, pois o F de significação é menor que o F crítico. Outra forma de analisar é por meio do p-valor, que neste caso, é menor que 5%, ou seja, a quantidade de partos tem relação com o número de leitões nascidos vivos. O valor do R quadrado foi de aproximadamente 87,83%, este valor mostra o quanto da variação da quantidade de leitões nascidos vivos é explicada pela quantidade de partos.

O modelo de regressão linear obtida está apresentado na Equação (2).

$$\text{quantidade de leitões nascidos vivos} = -25,52 + 12,86 * \text{quantidade de partos} \quad (2)$$

A regressão não é o modelo indicado para prever valores fora da variável explicativa, logo, utilizou-se para prever valores dentro da amostra, porém não existentes, como por exemplo, se houver um número de 32 partos, o modelo prevê 386 leitões nascidos vivos. Ou então, se houver 60 partos, prevê-se 746 leitões nascidos vivos.

Através da Equação (2) de regressão, foram previstos os valores da variável nascidos vivos (y), para os dados amostrais. Os resultados encontrados estão dispostos na Figura 4.

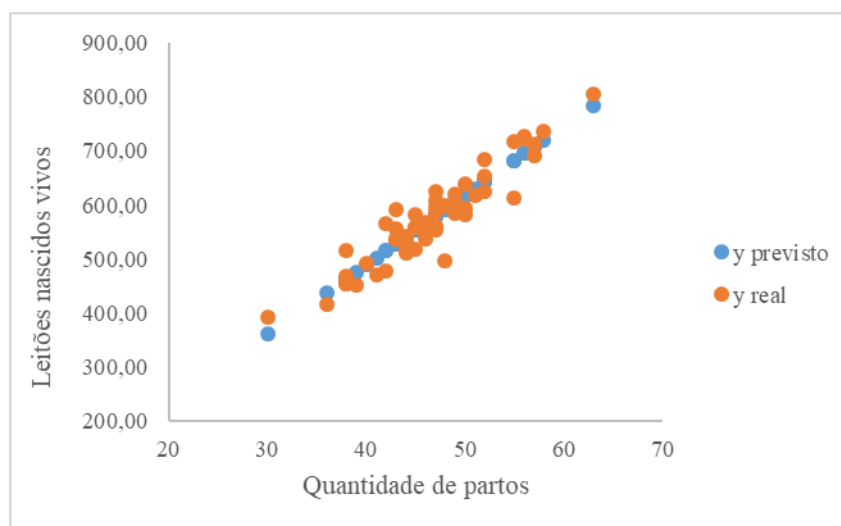


Figura 4 – Quantidade de partos X Comparação entre dados reais e previstos

Pela Figura 4, pode-se perceber a similaridade dos dados previstos e dos dados reais.

Por meio dos resultados desta regressão, foi feita a análise dos resíduos, que permite verificar se a dispersão dos dados em torno da reta estimada é muito grande.

A primeira análise foi em relação à linearidade que está disposta na Figura 5, esta análise diz que o modelo é adequado se, no gráfico dos resíduos *versus* variável dependente, não existir padrão de tendência.

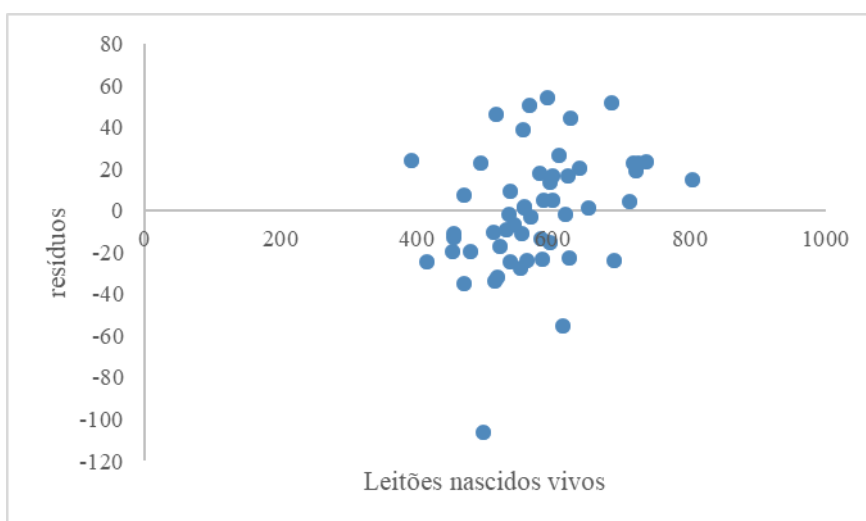


Figura 5 - Variável dependente X resíduos

Pela Figura 5, conclui-se que não há tendência nos dados dos resíduos que se apresentam dispersos, isso indica que o modelo está aceitável.

A segunda análise realizada nos resíduos foi da independência dos erros através do dos resíduos *versus* a ordem de coleta. Estes valores precisam estar dispersos no gráfico para o modelo ser aceitável. A Figura 6 mostra esta relação.

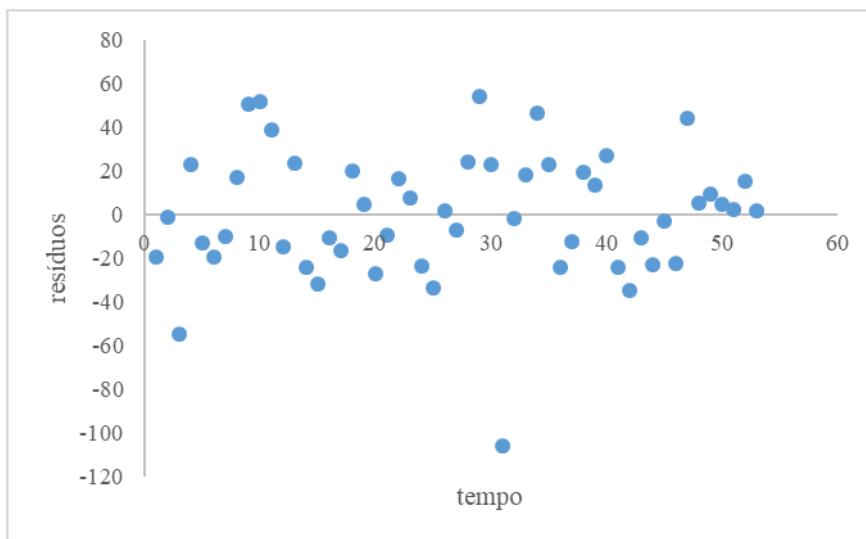


Figura 6 - tempo X resíduos

Pela Figura 6 pode-se perceber que os dados estão bastante dispersos, isso indica que o modelo está adequado.

No terceiro gráfico foi realizada a análise das variâncias, onde utiliza-se da variável independente e do resíduo. Para o pressuposto não ser violado, os dados não podem apresentar tendência. A Figura 7 mostra estes resultados.



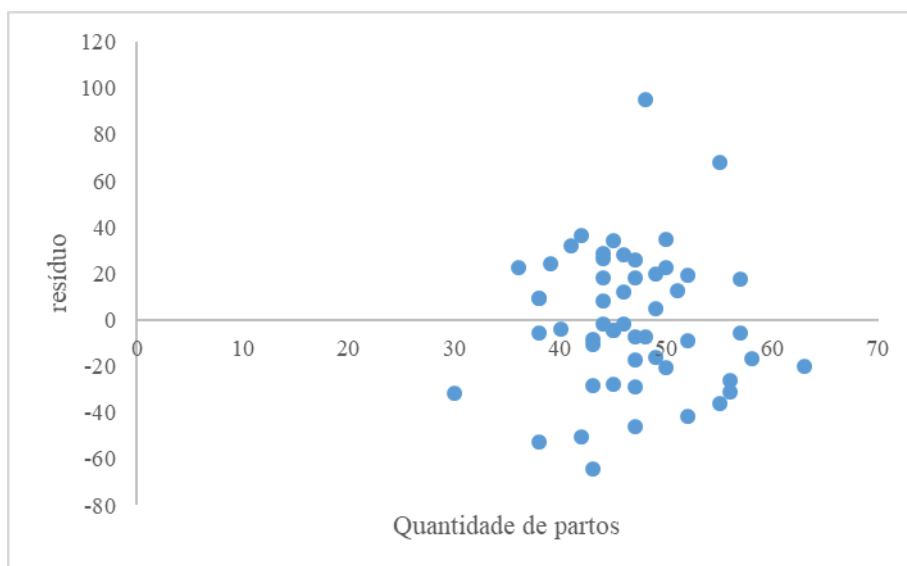


Figura 7 - variável independente X resíduos

Analisando a Figura 7, pode-se concluir que não existe tendência no gráfico de variância. Desta maneira, conclui-se que a regressão simples deste caso foi aceita pois atendeu aos três pressupostos da regressão linear.

A segunda análise foi realizada com relação a variável resposta nascidos vivos e a variável explicativa quantidade de coberturas. Para esta análise foram seguidos os mesmos passos da análise anterior. Os resultados podem ser visualizados na Tabela 2.

ESTATÍSTICA DA REGRESSÃO								
<b>R múltiplo</b>	0,01							
<b>R-quadrado</b>	9,71E-05							
<b>R-quadrado ajustado</b>	-0,02							
<b>Erro padrão</b>	87,06							
<b>Observações</b>	53							
	<i>GL</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>			
<b>Regressão</b>	1	37,52	37,52	0,0049	0,94			
<b>Resíduo</b>	51	386525,79	7578,93					
<b>Total</b>	52	386563,32						
	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
<b>Interseção</b>	562,71	162,46	3,46	0,001	236,55	888,87	236,55	888,87
<b>quantidade de coberturas</b>	0,21	2,93	0,07	0,94	-5,67	6,09	-5,67	6,09

Tabela 2 – Resultados da regressão considerando as variáveis de Quantidade de Coberturas e Nascidos Vivos

Na Tabela 2, observou-se que o valor do F de significação é maior que o F crítico, dessa maneira, não há relação significativa entre os dados. Notou-se que o resultado obtido para o p-valor foi superior a 0,05, reforçando a inexistência de relação. O valor do R-quadrado encontrado foi 0,009%, essa é a porcentagem de relação entre as duas variáveis analisadas, e o valor encontrado não satisfaz os requisitos, ao nível de significância adotado não apresenta uma relação linear da variável quantidade de coberturas e nascidos vivos.

A terceira análise teve como variável explicativa o número de marrãs cobertas e variável resposta o número de leitões nascidos vivos. Os resultados estão apresentados na Figura 10.

ESTATÍSTICA DA REGRESSÃO								
R múltiplo	0,05							
R-quadrado	0,002							
R-quadrado ajustado	-0,02							
Erro padrão	86,96							
Observações	53							
	<i>GL</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>			
Regressão	1	926,09	926,09	0,12	0,73			
Resíduo	51	385637,23	7561,51					
Total	52	386563,32						
	<i>Coeffi- entes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	567,23	23,02	24,64	5,4e-30	521,01	613,44	521,01	613,44
quantidade de coberturas	1,09	3,14	0,35	0,73	-5,21	7,41	-5,21	7,41

Tabela 3 – Resultados da regressão considerando as variáveis de Marrãs Cobertas e Nascidos Vivos

Da Tabela 3, observou-se que  $F=0,12$  que comparado com F de significação de 0,73, resultou em um valor inferior. Então, considera-se que não há uma relação linear entre as variáveis. Somado a isso, o p-valor resultou em um valor superior a 5%, afirmando que não existe relação entre as variáveis em questão. Já o valor do R-quadrado atestou a relação de apenas 2% entre essas variáveis.

Seguidamente, analisou-se a relação entre a variável estoque de matriz ativas e o número de leitões nascidos vivos, e os resultados estão apresentados na Figura 11.

ESTATÍSTICA DA REGRESSÃO								
<b>R múltiplo</b>	0,012							
<b>R-quadrado</b>	0,00014							
<b>R-quadrado ajustado</b>	-0,019							
<b>Erro padrão</b>	87,05							
<b>Observações</b>	53							
	<i>GL</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>			
<b>Regressão</b>	1	56,37	56,37	0,007	0,93			
<b>Resíduo</b>	51	386506,94	7578,56					
<b>Total</b>	52	386563,32						
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
<b>Interseção quantidade de coberturas</b>	610,3	419,72	1,45	0,15	-232,32	1452,92	-232,32	1452,92
	-0,04	0,41	-0,09	0,93	-0,86	0,79	-0,86	0,79

Tabela 4 – Resultados da regressão considerando as variáveis Estoque de Matrizes Ativas e Nascidos Vivos

Na Tabela 4, foi observado que o valor de F de significação é superior ao valor F, o que mostra que a relação não é significativa, com o p-valor resultando em 0,93 comprovou a análise anterior pelo fato de ser maior que 0,05. E por fim a relação de 0,01% obtido no R-quadrado reafirmou a não influencia do estoque de matriz ativas nos nascidos vivos.

Por meio dos resultados das análises dentre todas as variáveis explicativas, conclui-se que somente a variável quantidade de partos irá influenciar linearmente a quantidade de leitões nascidos vivos, sendo possível realizar previsões através do modelo obtido e apresentado na Equação (2).

## Conclusões

Este artigo objetivou prever a quantidade de leitões nascidos vivos em uma unidade produtora de leitões localizada na cidade de São Miguel do Iguaçu. Com o alto crescimento da suinocultura torna-se importante investir no gerenciamento de granjas suínas. A previsão vem como ferramenta importante para auxiliar o processo de tomada de decisão sobre investimentos, consumo de matérias primas, entre outros. Deste modo, ajudando a evoluir a cadeia produtiva de suínos no Oeste do Paraná.

## Referências

Associação Brasileira de Proteína. **Relatório Anual**. 2017. Disponível em: <<http://abpa-br.com.br/setores/avicultura/publicacoes/relatorios-anuais>>. Acesso em: 12 set. 2017.

**ATIVOS SUINOCULTURA**. CNA Brasil: CNA Brasil e CEPEA, 2015. Disponível em: <<http://www.cnabrazil.org.br/boletins/ativos-suinocultura-suinocultura-brasileira-avanca-no-cenario-mundial-maio-2015>>. Acesso em: 17 mai. 2017.

BARBETTA, Pedro Alberto; REIS, Marcelo Antônio; BORNIA, Antônio Cezar. **Estatística para cursos de engenharia e informática**. 3. ed. São Paulo: Editora Atlas, 2010. 410 p.

CORRAR, Luiz J.; PAULO, Edilson; FILHO, José Maria. **Análise Multivariada**. 1.ed. São Paulo: Atlas S.A., 2012.

FERNANDES, Flavio Cesar Faria; GODINHO FILHO, Moacir. **Planejamento e controle da produção: dos fundamentos ao essencial**. 2. ed. São Paulo: Editora Atlas, 2010. 275 p.

KRAJEWSKI, Lee; RITZMAN, Larry; MALHOTRA, Manoj. **Administração de produção e operações**. 8. ed. São Paulo: Pearson, 2014. 615 p.

MARTINS, Petrônio G.; LAUGENI, Fernando Piero. **Administração da produção**. 2. ed. São Paulo: Editora Saraiva, 2013. 562 p.

MONKS, Joseph G. **Administração da produção**. São Paulo: Editora Mcgraw-hill Ltda, 1987. 502 p.

ROHR, Stefan Alexander; COSTA, Osmar Antônio dalla; COSTA, Filipe Antônio dalla. **Bem-estar na produção de suínos: Toda granja**. Brasília: EMBRAPA e ABCS, 2016. 38 p.

TUBINO, D. F., **Planejamento e controle da produção**. 2 ed. São Paulo: Atlas, 2007. 190 p.